AE

Eryurtlu 2

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number : **0 577 363 A1**

(12) **EUROPEAN PATENT APPLICATION**

(21) Application number : **93305022.1**

(22) Date of filing : **28.06.93**

(51) Int. Cl.5 : **H04N 7/13**, H04N 7/133

(30) Priority : **30.06.92 US 906787**

(43) Date of publication of application : **05.01.94 Bulletin 94/01**

(84) Designated Contracting States : **DE FR GB IT**

(71) Applicant : **LORAL AEROSPACE CORPORATION**
600 Third Avenue
New York, NY 10016 (US)

(72) Inventor : **Wilson, Dennis L.**
1325 Breenwood Avenue
Palo Alto, CA 94301 (US)
Inventor : **Stenstrom, James T.**
877 Heatherstone Way
518 Mountain View, CA 94040 (US)

(74) Representative : **Vaufrouard, John Charles**
Elkington and Fife Prospect House 8
Pembroke Road
Sevenoaks, Kent TN13 1XR (GB)

(54) **Compression and reconstruction of radiological images.**

(57) The present invention provides for a compression scheme (10) tailored to the compression of large radiological images processed with image processing workstations. The compression scheme (10) is a variation of the CCITT JPEG compression scheme with special care taken to suppress blocking effects of the 8x8 discrete cosine transform used therein. In accordance with the present invention, an original image (11) is downsampled (12), then interpolated (14) back to its original size, resulting in a smoothed image (14a). The difference (13) between the smoothed image (14a) and the original image (11) is then compressed (15) using an 8x8 discrete cosine transform with 12 bit data. Major artifacts present in the original image (11) are due to the mismatching of the low frequency components of the image at the edges of the blocks. When the low frequencies are removed by subtracting out the low frequency components, the artifacts disappear except in areas having very large dynamic changes. The compression scheme (10) of the present invention is particularly useful for archiving or telecommunicating images, achieving compression of more than 10:1 without losing the characteristics of the radiological image that permit effective diagnosis. A reconstruction scheme is also disclosed which allows recovery of the image (11) from a compressed image file.

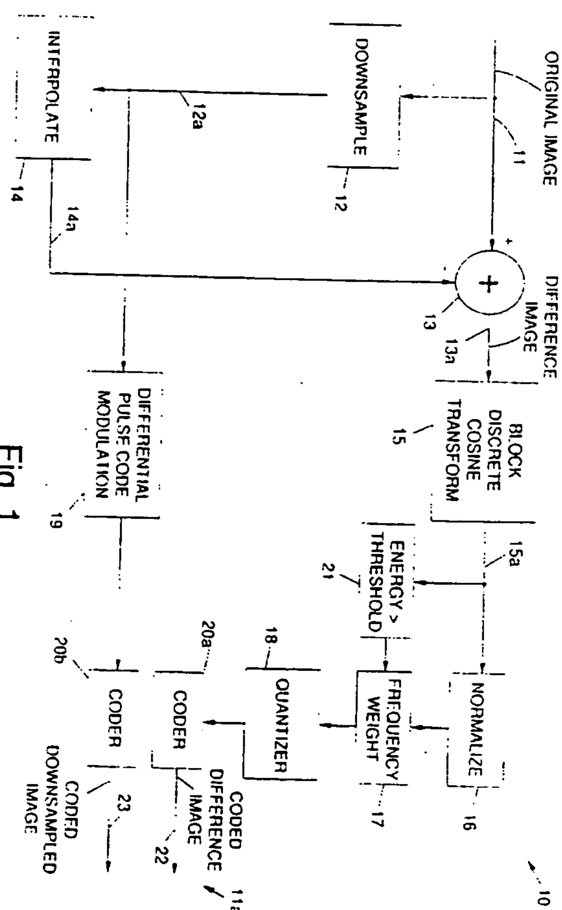


Fig. 1

EP 0 577 363 A1

BACKGROUND

The present invention relates generally to radiological imaging, and more particularly, to image compression apparatus and methods for use in compressing and reconstructing radiological images.

The medical imaging industry is a moving away from traditional film-based systems and is moving toward the use of digital images presented on computer work-stations. The images that are presented to radiologists must preserve the information contained in the original film images in order for soft copy images to be accepted by the medical community. The typical radiological image is different from other images in several ways.

Radiological images used in medicine present peculiar problems for compression compared to conventional image compression problems. The images locally have very low contrast. For example, a transition over an edge of a bone does not change the image intensity by a large amount. While the local variations are small, the dynamic range over the whole image may be quite large. Much of the information is contained in the local variation of image intensity. The overall image provides a context within which the information is embedded, but the pathology is visible as local variations. The images are subject to large changes in the individual pixel intensity as the radiologist views the image. A radiologist may change the contrast and the center of the contrast range in his efforts to see details in the image. In the extreme, the radiologist may invert the image, changing white to black. The images are very large, typically more than 1k by 1.2k pixels and up to 2k by 2.5k pixels. The image dynamic range typically requires 12 bits. The image is viewed at a number of magnifications from 2:1 minified, to 8:1 zoomed.

The well-known CCITT JPEG standard uses an 8x8 discrete cosine transform in the image compression scheme. The number of bits in the image is 8 or 12 bits. The blocking effect of the discrete cosine transform is unacceptable as an artifact over the entire image. The dynamic range of 12 bits cannot be squeezed into 8 bits. If a typical JPEG compressed image is examined in detail, the blocking of the 8x8 discrete cosine transform is visible over the whole of the image. The details of the image are visible through the blocking, but the distraction of the blocking over the whole of the image is unacceptable. The radiologist often looks at a zoomed image with contrast arranged so that the blocks interfere with his viewing of the image.

If radiological images were viewed in the same way that a TV image is viewed, the dynamic range could be compressed to 8 bits by a simple nonlinear mapping of image intensities. A typical mapping scheme is logarithmic mapping. With this scheme, displayed image intensities are proportional to the logarithm of the original image intensities. Since the image contrast and intensity center can be varied, including inverting the pixel values, such nonlinear mapping is not appropriate. No companding of the image intensity values is tolerable, since the compression from 12 bits to 8 bits introduces large artifacts at some level of contrast or center for the displayed intensity values.

Other techniques that are used for image compression are typically full image compression schemes. Either a Fourier transform or a cosine transform is used over the full image. Since the transform is a full image transform, there is no blocking of the image. The difficulty with the full image transform is that the amount of computation per pixel is large and the images, in many cases, are constrained to be square images. If the images are not square, they are padded to make them square. In many implementations, the number of pixels in the image that is compressed is constrained to be a power of two. Again the images must be padded with zeros to make them a power of two.

In view of the above, it is an objective of the present invention to provide for apparatus and methods that provide for image compression of radiological images that suppresses artifacts present in the image due to blocking.

SUMMARY OF THE INVENTION

The present invention provides for a compression scheme tailored to the compression of large radiological images displayed using image workstations. The present invention may be implemented as a special purpose processor for use in a computer workstation, for example, or may be implemented in firmware or software if so desired. The compression scheme is a variation of CCITT JPEG compression that is adapted to suppress blocking effects caused by the processing of images with discrete cosine transforms used in conventional JPEG compression. The compression is particularly useful for archiving or telecommunicating images, achieving compression of more than 10:1 without losing the characteristics of the original radiological image. Furthermore, the apparatus and method of the present invention produces a compressed radiological image that permits effective diagnosis.

In order to achieve the above and other objectives, the present invention provides for image compression apparatus and methods for use with radiological images that suppresses artifacts present in the image due to blocking. The present invention reduces the size of digitized radiological images for storage or for transmission such that the images remain clinically useful for diagnosis. The present invention provides for the separation of low frequency components for suppression of compression artifacts

combined with a block mode transform compression for localizing compression noise. The present invention suppresses block artifacts in compression while providing for compression noise with a controlled local signal-to-noise ratio.

The suppression of artifacts due to blocking is accomplished by the use of a difference image. The original image is downsampled, then interpolated back to its original size to produce a smoothed image. The difference between the smoothed image and the original image is then compressed using an 8x8 discrete cosine transform with 12 bit data. The major artifacts present in the original image are due to the mismatching of the low frequency components of the image at the edges of the blocks. Consequently, when the low frequencies are removed by subtracting out the low frequency components, the artifacts disappear except in areas having very large dynamic changes.

To make the quantization noise relative to the magnitude of local variations in the image, the total energy in the transform block is normalized, then weighted, quantized, and coded. The normalization is accomplished by finding the sum of the squares of coefficients of the cosine transform. A gain factor is determined such that the energy in a block is equal to a constant. This gain factor is applied to each of the coefficients of the transform, changing the energy in the block to a constant value that is equal to the normalization energy. The coefficients are then quantized by rounding (not truncating) to the nearest integer. An exception is made if the energy in the block is larger than a predetermined threshold. In this case, no normalization is performed. The coefficients are quantized at the original level. When the energy is high, the dynamic range of the image in that vicinity is high, leading to more prominent blocking effects. When the block is not normalized, the quantization noise is reduced, thus suppressing the blocking effects.

The coding of the weighted coefficient values may be done in several ways. The use of a Huffman code or an approximate Huffman code is one technique. The Huffman code uses short code words to transmit coefficient values that happen often and longer code words to transmit coefficient values that are less probable. The average code word length is smaller using this approach than using a fixed length code word. An energy normalization parameter controls the amount of compression accomplished by the Huffman coding. When the normalization constant is small, the numbers to be encoded are small. The reduction of the range of numbers to be encoded reduces the number of bits required to encode the numbers.

An alternative to Huffman coding is to use arithmetic coding. This approach codes the weighted, quantized coefficients in a code word that maps the coefficient values into the interval from zero to one. The arithmetic coding scheme may be made adaptive, and thus be independent of the type of image in contrast to the image type sensitivity of the Huffman coder.

The discrete cosine transform yields a set of transform coefficients. The coefficients roughly describe the energy levels at frequencies in the row and column directions in the image. The human eye is less sensitive to noise at high frequencies, thus permitting coarser quantization at higher frequencies. The coarser quantization is accomplished by frequency weighting the transform coefficients. Lower weights are given to higher frequency components.

The compression of the image is accomplished by deleting the high frequency components with little energy and using Huffman (or other) coding on the remaining quantized discrete cosine transform coefficients to compress them. The level of quantization of the discrete cosine transform coefficients is adjusted by setting a quantization level that depends on the energy in the block. When the block energy is high, the quantization level is high. With such large images the noise generated by the quantization is a white noise over the entire image. The human eye can look through such noise, averaging the intensity values in a local region to suppress the noise. Even severe noise only appears to be a general haziness to the image.

BRIEF DESCRIPTION OF THE DRAWINGS

The various features and advantages of the present invention may be more readily understood with reference to the following detailed description taken in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

Fig. 1 shows an image compression apparatus and method in accordance with the principles of the present invention;

Fig. 2 shows how high frequency component deletion is performed in the apparatus and method of Fig. 1 starting at the highest frequency and proceeding toward lower frequencies until a desired percent of the energy has been removed;

Fig. 3 shows a differential pulse code modulation encoding method employed in the apparatus and method of Fig. 1;

Fig. 4 shows an image reconstruction apparatus and method in accordance with the principles of the present invention, which uses a minified image and a recovered difference image to reconstruct the original image; and

Fig. 5 shows recovery of the downsampled image in the image reconstruction apparatus and method of Fig. 4.

DETAILED DESCRIPTION

Referring to the drawing figures, the architecture and processing flow of an image compression apparatus and method 10 in accordance with the principles of the present invention is shown in Fig. 1. More particularly, Fig. 1 shows the use of downsampling and interpolation processes that eliminate low frequency components from a discrete cosine transformed image. The suppression of artifacts in an original image 11 due to the blocking is accomplished by the use of a difference image 13a. An original image 11 is downsampled 12 to produce a downsampled image 12a, which is then interpolated 14 back to the size of the original image 11. Interpolating the downsampled image 12a produces a smoothed image 14a. The difference image 13a is produced between the original image 11 and the interpolated image 14a using an adder 13. The difference between the smoothed image 14a and the original image 11 is then compressed 15 using an 8x8 pixel block discrete cosine transform with 12 bit data. This produces a transformed image 15a. Major artifacts that are present in the original image 11 are due to the mismatching of low frequency components in the original image 11 at the edges of the blocks. When the low frequencies are removed and separately coded 20a by subtracting out the low frequency components to produce a coded difference image 22, the artifacts disappear except in areas of very large dynamic changes.

To make quantization noise present in the difference image 13a relative to the magnitude of local variations, the total energy in the transform block (transformed image 15a) produced by compression is normalized 16, and is then weighted 17, quantized 18, and encoded by the coder 20a. The normalization 16 is accomplished by finding the sum of the squares of the coefficients of the discrete cosine transform. A gain value such that the block energy is equal to a constant is determined. This gain value is applied to each of the coefficients of the transform, changing the energy in the block to a constant value equal to the normalization energy. The coefficients are then quantized 17 by rounding (not truncating) them to the nearest integer. An exception is made if the energy in the block is larger than a predetermined threshold. In this case, no normalization 16 is performed, and the coefficients are quantized 17 at the original level. When the energy is high, the dynamic range of the image in that vicinity is high, leading to more prominent blocking effects. When the block is not normalized 16, the quantization noise is reduced, suppressing the blocking effects. The above processing produces the coded difference image 22 that is a first part of a compressed image file 11a.

The downsampled image 12a is differential pulse code modulated 19 and is also encoded in a second coder 20b using the same coding scheme as is used for the quantized, weighted transform coefficients. This produces a coded downsampled image that forms a second part of the compressed image file 11a.

The encoding 20a of the weighted coefficients may be done in several ways. The use of a Huffman code or an approximate Huffman code is a technique to code the weighted coefficients. The well-known Huffman code uses short code words to transmit coefficients that happen often and longer code words to transmit coefficients that are less probable. The average code word length is smaller using this approach than using a fixed length code word. An energy normalization parameter controls the amount of compression 15 accomplished by the Huffman coding. When the normalization constant is small, the numbers to be encoded are small. The reduction of the range of numbers to be encoded reduces the number of bits required to encode the numbers.

An alternative to the use of Huffman coding is the use of arithmetic coding. This approach codes the weighted, quantized coefficients in a code word that maps the coefficients into an interval from zero to one. This scheme may be made adaptive, and is thus independent of the type of image in contrast to the image type sensitivity of the Huffman coder.

The discrete cosine transform compression 15 produces a set of transform coefficients. The coefficients roughly describe the energy levels at frequencies in the row and column directions of pixels in the original image 11. The human eye is less sensitive to noise at high frequencies, permitting coarser quantization 17 at the higher frequencies. The coarser quantization 17 is accomplished by frequency weighting the transform coefficients. Lower weights are given to the higher frequency components.

The compression achieved by the present invention is accomplished in two ways, the deletion of the high frequency components with little energy and the use of Huffman coding, for example, on the remaining quantized discrete cosine transform coefficients. The level of quantization 17 of the discrete cosine transform coefficients is adjusted by setting a quantization level that depends on the energy in the block. When the block energy is high, the quantization level is high. With such large images the noise generated by the quantization 17 is a white noise over the entire original image 11. The eye looks through such noise, averaging the intensity values in a local region to suppress the noise. Even severe noise appears to be only a general haziness to the image.

On the other hand the deletion of the high frequency components of the original image 11 tends to make the image 11 less crisp. Carried to an extreme, essential detail begins to disappear from the image 11. If the energy were uniformly distributed over the discrete cosine transform coefficients, the deletion of one coefficient removes about 1.6% of the energy.

Ninety percent of the energy would be contained in 57 of 64 coefficients. The energy is not uniformly distributed and is contained mostly in the low frequency components. Typically 45 coefficients remain when the energy retained is 95% of the original energy and 12 coefficients are retained when the energy retained is 50% of the energy.

The implementation of the elements of the compression scheme of the present invention are important to successful suppression of the blocking. In particular the filtering used in downsampling 12 and interpolation 14 back to the original image size must be done carefully in order to remove a large part of the low frequency energy of the original image 11 and to form a seamless connection across the blocks. A suitable filter is a raised cosine or a Kaiser window with a span of 9 pixels. The same filter is used for downsampling 12 and for interpolation 14 to restore the original image size. The span of 9 pixels means that for these two filter types, the weighting for the pixel that is separated from the center of the filter by four pixels is zero. The result is that the filter actually requires seven multiply-adds to form a point in the downsampled image 14a.

The downsampling 12 is performed in two steps, downsampling in the row dimension followed by downsampling in the column direction of the image 11. While a composite filter that performs the complete downsample operation in one step could be performed, the number of multiply-adds required is greater than the one dimension at a time approach.

The interpolation 14 is performed in the row and column direction simultaneously. The interpolation 14 may be viewed as inserting pixels in the downsampled image 12a by adding pixels with zero value followed by a filtering operation. The filter is the above-described two dimensional raised cosine or Kaiser window. The actual number of multply-adds required to build the image is limited, since most of the pixel values are zero. For some of the pixels no multiply adds are required, since there is only one non-zero pixel and the weight for that pixel is one. For other pixels two multiply-adds are required, while for yet others, four multiply adds are required. The organization of the sequence of multiply-adds so that only the required computations are performed makes for fast operation.

Normalization 16 proceeds by forming the square-root of the sum-of-squares of the discrete cosine transform coefficients. A multiplier that normalizes the energy of the block of discrete cosine transform coefficients is found by dividing the normalized energy constant, a parameter of the compression, by the energy in the block. Each of the coefficient values is multiplied by the resulting value.

An exception is made if the energy in the block is larger than a predetermined threshold value, as is indicated by the bypass path around the normalization step 16a and through a threshold determination step 21 to the weighting step 16. This large value of the energy in a block of the difference image 13a indicates that there is an area with high dynamic range within the block. This high dynamic range shows the effects of the quantization and the blocking more prominently than does the surrounding lower energy blocks. In these areas no normalization is performed, as is indicated by the bypass path through the threshold determination step 21. The discrete cosine transform coefficients are quantized 18 by rounding to form integers without modifying the amplitude of the coefficients.

The frequency weighting 16 is performed after normalization 17 and before quantization 18 of the coefficients. The frequency weight is, for example, a simple exponential weighting function: $W(\text{row},\text{column}) = e^{(-(\text{row}+\text{column})/c)}$, where W is the weight, row is the number of the row in the block from 0 to 7, column is the number of the column within the block from 0 to 7, and c is the weighting parameter. After the frequency weighting 16, the discrete cosine transform coefficients are quantized 18 by rounding them to form integers as was discussed above.

With reference to Fig. 2, the high frequency values are removed by deleting the high frequency components in order (illustrated by the arrowed line) until a percentage of the energy in the block has been removed. The percentage of the energy that is retained is a parameter of the compression. By including the number of coefficient values that have been retained in the coded block information, the level of deletion is retained.

Referring to Fig. 4, on reconstruction, the coded downsampled image 23 is reconstructed by decoding the Huffman (or other) code and inverting the differential pulse code modulation process 18. The downsampled image 23 is interpolated back to the full image size in the same manner as used in the compression process. The interpolated image is ready to be added to the difference image as soon as it is reconstructed.

The reconstruction of the coded difference image 22 to produce a recovered difference image 26a is performed by decoding the Huffman (or other) code recovering the weighted, normalized discrete cosine transform coefficients. The weighting and normalization are removed 24 by multiplying by the reciprocal of the weights used in during compression, and by multiplying each coefficient by the ratio between the normalization energy and the energy in the block. The energy in the block is part of the coded information with the coefficients. The normalization energy is a parameter of the compression and is passed with the compressed image to permit decompression.

The difference image is recovered by performing an inverse discrete cosine transform 26 on the cosine transform coefficients. The original image 11 is recov-

ered by adding the recovered difference image 26a to a recovered interpolated downsampled image 29a.

Expansion of the compressed image 11a on recovery from archive is as follows. The expansion of the image follows the inverse of the compression steps. The expansion architecture and processing steps are shown in Fig. 4. The minified image (coded downsampled image 23) and the normalization value are recovered. The minified image (coded downsampled image 23) is recovered by performing a bit-preserving reconstruction of the image as is shown in Fig. 5. The coded difference image 22 is recovered by recovering the quantized discrete cosine transform coefficients using the block energy value recorded is a second word of a compressed block. The inverse discrete cosine transform 26 is then performed, resulting in the recovered difference image 26a. The minified image (coded downsampled image 23) is then interpolated 29 back to full size and added to the recovered difference image 26a in the adder 13 to recover the approximation to the original image 11 (reconstructed original image 11b).

The differential pulse code modulation processing 28 of the coded downsampled image 23 is shown in more detail in Fig. 5. The simplicity of the differential pulse code modulation encoder is maintained through the use of the differential pulse code modulation decoder 28. The result of using the differential pulse code modulation decoder 28 is the recovered downsampled image 28a that is to be interpolated 29 back to the full size (recovered interpolated image 29a) and added 30 to the recovered difference image 26a derived from the cosine transform processing 26.

The use of the compression scheme of the present invention results in a noise value that is the difference between the original image 11 and the reconstructed original image 11b. The use of the energy threshold frequency cutoff and the energy normalization level establish the noise level relative to the image energy level. Since the low frequencies are removed in the minified image 23, the energy that is considered is the energy of the high frequencies. This removal of the low frequencies before compression makes the compression local. That is, the compression depends not on the image intensity over a large area, but on the changes in intensity over a small area. Considering the noise to be the departure from the difference image 13a, the signal-to-noise ratio can be calculated, and is given by:

$SNR_{deletion}$ = 10 log(1 - proportion_energy_deleted)

A second component of the signal-to-noise ratio is the quantization noise. If the quantization level is established with reference to the energy level of the block, the quantization noise component can be established by analysis, and wherein

$SNR_{quantization}$ = 10.79 + 20 log(Normalization_level), where Normalization_level is the level that the signal energy is raised to before the quantization 18 is performed. The quantization 18 is done by rounding, not by truncation. Truncation loses the 10.79 dB of signal-to-noise ratio.

The two components of noise add as the sum of squares. The composite signal-to-noise ratio is therefore the square root of the sum of the squares of the individual noise values compared to the energy in the difference image 13a. By this analysis, the present invention is balanced when the two noise sources generate equal amounts of noise. In fact, the eye is more sensitive to the loss of high frequencies than to the added general noise in this application. It is therefore appropriate to make the quantization noise about 3 times the noise due to the deletion of the high frequency energy.

The use of the weighting coefficient on compression requires the use of a reciprocal weighting coefficient on the expansion in order to restore the original values to the image 11. It is possible for the values of the weighting coefficient to be negative and positive. In this case, the high frequencies are accented on compression, reducing the amount of compression and increasing the quality of the compressed image.

The compressed image 11a has two segments, the minified image 23 appropriately coded and the discrete cosine transform compressed coded difference image 22. The format of the file of the compressed image 11a is the concatenation of a small image plus the compressed large image. Each block of the large image consists of a 6 bit word indicating the number of coefficients retained, a word indicating the energy level of the block in the coded difference image 22, and the remaining discrete cosine transform coefficients. The block is Huffman coded to support the last step of the compression. In the image file header, there must be an indication of the level of the normalization of the image block energy.

The small image is encoded using a simple differential pulse code modulation scheme shown in Fig. 3. Successive differences in the pixel values are encoded using the Huffman encoder. The use of the something more than the very simple encoding is not warranted. There are 1/16 the number of pixels in the small image compared to the large image. The use of the Huffman coder reduces the number of bits by a factor of approximately 2.5 depending on the image. The result is a compression ratio of 40:1 for the small image. Compression of the difference image 13a by a factor of 40:1 results in an overall compression ratio of 20:1, an acceptably high compression ratio with very little degradation. Further compression of the difference image 13a results in unacceptable degradation. It is therefore of little value to further reduce the size of the compressed small image. However, the further reduction of the minified image 23 may be accomplished by applying minification along with discrete cosine transform compression of the difference image 13a.

The perceived image quality provided by the present invention is very good. The compression from 12 bits per pixel to about 1.2 bits per pixel leave an image that is very difficult to discriminate from the original image 11. Beyond this factor of 10:1 compression the image begins to degrade slightly. At 15:1 compression, high contrast areas of the image begin to show the blocking of the discrete cosine transform. At 20:1 compression sharp edges of the image are noticeably softer due to the loss of the high frequencies.

There are three compression parameters, the energy normalization constant, the percent energy retained, and the frequency weighting coefficient. The energy normalization constant is a number between 5 and 5000. The percent energy retained is a number between 1 and 100. The frequency weighting coefficient is a number between 2 and 10. The compression may be controlled using a single number between 1 and 100. The compression control number can be mapped into the three compression parameters using a suitable function to establish the quality of compression.

The amount of compression is variable. If there is much detail in the image, the compressed image is longer than an image with little detail.

Thus there has been described a new and improved image compression apparatus and methods for use in compressing and reconstructing radiological images. It is to be understood that the above-described embodiment is merely illustrative of some of the many specific embodiments which represent applications of the principles of the present invention. Clearly, numerous and other arrangements can be readily devised by those skilled in the art without departing from the scope of the invention.

Claims

1. An image compression method (10) characterized by the steps of:

   downsampling (12) an original radiological image (11) to produce a downsampled image (12a);

   interpolating (14) the downsampled image (12a) back to its original size to provide a smoothed image (14a);

   processing (13) the smoothed image (14a) and the original image (11) to provide a difference image (13a);

   processing (15) the difference image (13a) using a discrete cosine transform to provide a transformed difference image (15a);

   frequency weighting (17) the transformed difference image (15a) to provide a weighted difference image;

   quantizing (18) the weighted difference image to provide a quantized difference image;

   encoding (20a) the quantized difference image to produce an encoded difference image (22);

   differential pulse code modulating (19) the downsampled image to produce a differential pulse code modulated downsampled image; and

   encoding (20b) the differential pulse code modulated downsampled image to produce a coded downsampled image (23);

   whereby the encoded difference image and the encoded downsampled image comprise a compressed image, and wherein the compression is adapted to remove artifacts present in the original radiological image caused by blocking.

2. The method (10) of Claim 1 further comprising the step of normalizing (16) the compressed image prior to the frequency weightng (17), characterized by the steps of:

   finding the sum of the squares of the coefficients of the discrete cosine transform;

   determining a gain factor such that energy in the transformed difference image is equal to a constant; and

   applying the gain factor to each of the coefficients of the discrete cosine transform, changing the energy in the transformed difference image to a constant value that is equal to the normalization energy.

3. The method (10) of Claim 1 wherein the step of quantizing (18) the compressed image is characterized by the step of rounding coefficients to quantize them to a nearest integer.

4. The method (10) of Claim 1 wherein, if the energy the transformed difference image is larger that a predetermined threshold, the step of quantizing (18) the compressed image comprises quantizing the coefficients at their original level so that the quantization noise is reduced, thus suppressing blocking effects.

5. The method (10) of Claim 1 wherein the step of encoding (20a) the quantized image is characterized by the step of encoding of the weighted coefficient values using a Huffman code.

6. The method (10) of Claim 1 wherein the step of encoding (20a) the quantized image is characterized by the step of encoding of the weighted coefficient values using an arithmetic code that codes the weighted, quantized coefficients in a code word that maps coefficient values into an interval from zero to one.

7. The method (10) of Claim 1 wherein the step of processing the difference image (13a) using a

discrete cosine transform is characterized by the step of frequency weighting the transform coefficients wherein lower weights are given to higher frequency components.

8. The method (10) of Claim 1 wherein the step of processing the difference image (13a) is characterized by the steps of:

deleting high frequency components having little energy; and

Huffman coding the remaining quantized discrete cosine transform coefficients.

9. The method (10) of Claim 8 wherein the level of quantization of the discrete cosine transform coefficients is adjusted by setting a quantization level that is a function of the energy in the transformed difference image.

10. An image reconstruction method (10) for use with compressed images comprising an encoded difference radiological image and an encoded downsampled radiological image that produces a reconstructed image, said method (10) characterized by the steps of:

decoding (24) the encoded difference image (22) to produce a decoded image (24a);

frequency weighting and removing the normalization of the decoded difference image (25) to produce a restored transformed difference image;

transforming (26) the restored transformed difference image using an inverse discrete cosine transform to produce a restored difference image (26a);

decoding (27) the coded downsampled radiological image (23) to produce a decoded image (27a);

differential pulse code demodulating (28) the decoded image (27a) to produce a demodulated image (28a);

interpolating (29) the differential pulse code modulated image (28a) to provide a full scale smoothed image (29a); and

adding (30) the compressed image (26a) and the smoothed image (29a) to produce a reconstructed original image (11b).

11. Image compression apparatus (10) characterized by:

means for downsampling (12) an original radiological image to produce a downsampled image;

means for interpolating (14) the downsampled image back to its original size to provide a smoothed image;

means for processing (13) the smoothed image and the original image to provide a difference image;

means for processing (15) the difference image using a discrete cosine transform to provide a transformed difference image;

means for frequency weighting (17) the transformed difference image to provide a weighted difference image;

means for quantizing (18) the weighted difference image to provide a quantized difference image;

means for encoding (20a) the quantized difference image to produce an encoded difference image;

means for differential pulse code modulating (19) the downsampled image to produce a differential pulse code modulated downsampled image; and

means for encoding (20b) the differential pulse code modulated downsampled image to produce a coded downsampled image;

whereby the encoded difference image and the encoded downsampled image comprise a compressed image, and wherein the compression is adapted to remove artifacts present in the original radiological image caused by blocking.

12. The apparatus (10) of Claim 11 further characterized by means for normalizing (16) the compressed image is characterized by:

means for finding the sum of the squares of the coefficients of the discrete cosine transform;

means for determining a gain factor such that energy in the transformed difference image is equal to a constant; and

means for applying the gain factor to each of the coefficients of the discrete cosine transform, changing the energy in the transformed difference image to a constant value that is equal to the normalization energy.

13. The apparatus (10) of Claim 11 wherein the means for quantizing (18) the compressed image is characterized by means for rounding coefficients to quantize them to a nearest integer.

14. The apparatus (10) of Claim 11 wherein, if the energy in the transformed difference image block is larger that a predetermined threshold, the means for quantizing (18) the compressed image is characterized by means for quantizing (18) the coefficients at their original level so that the quantization noise is reduced, thus suppressing blocking effects.

15. The apparatus (10) of Claim 11 wherein the means for encoding (20a) the quantized image is characterized by means for encoding of the

weighted coefficient values using a Huffman code.

16. The apparatus (10) of Claim 11 wherein the means for encoding (20a) the quantized image is characterized by means for encoding of the weighted coefficient values using an arithmetic code that codes the weighted, quantized coefficients in a code word that maps coefficient values into an interval from zero to one.

17. The apparatus (10) of Claim 11 wherein the means for processing (15) the difference image using a discrete cosine transform is characterized by means for frequency weighting (17) the transform coefficients wherein lower weights are given to higher frequency components.

18. The apparatus (10) of Claim 11 wherein the means for processing (15) the difference image is characterized by:

means for deleting high frequency components having little energy; and

means for Huffman coding the remaining quantized discrete cosine transform coefficients.

19. The apparatus (10) of Claim 18 wherein the level of quantization of the discrete cosine transform coefficients is adjusted by setting a quantization level that is a function of the energy in the transformed difference image.

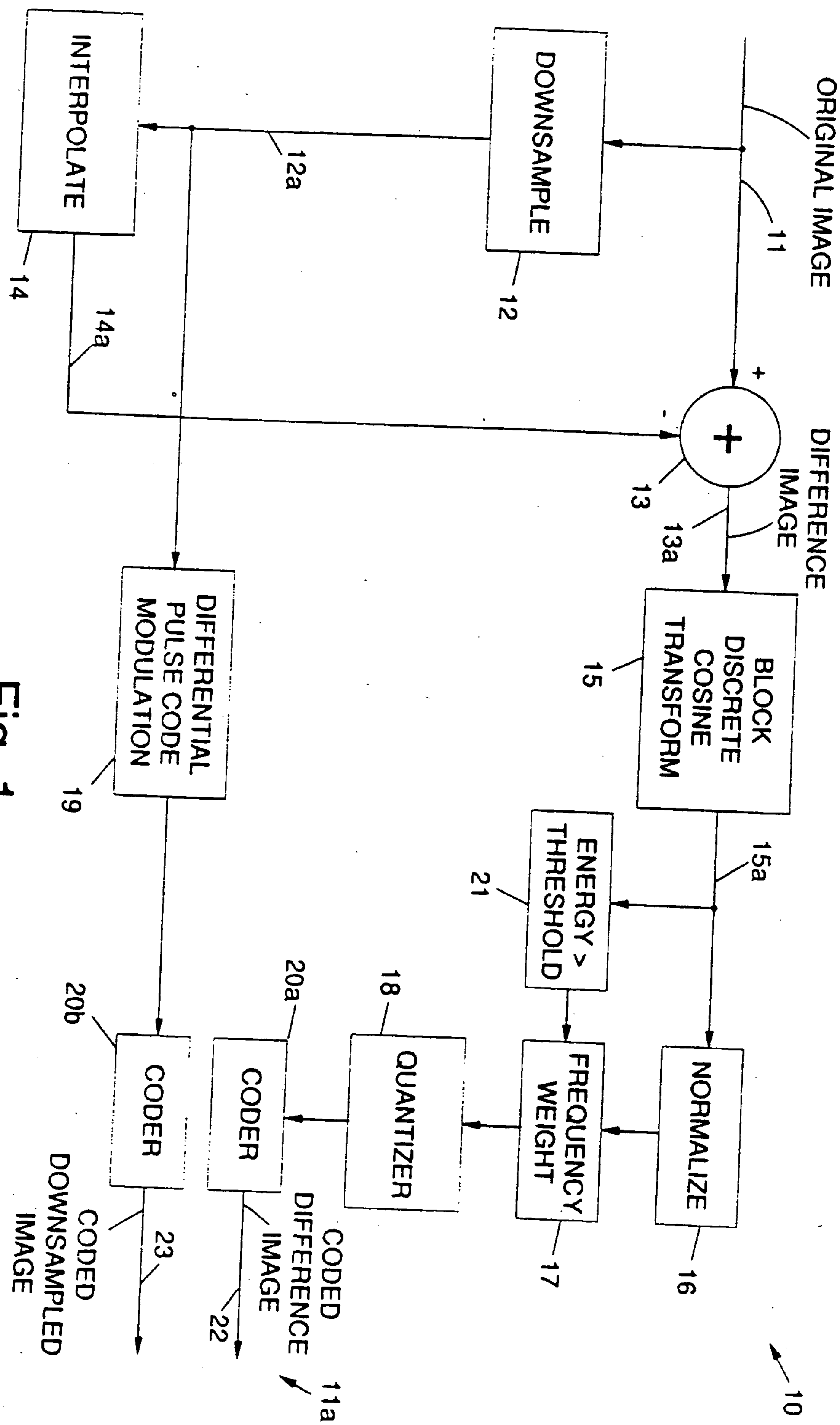

ORIGINAL IMAGE
11
DOWNSAMPLE
12
12a
INTERPOLATE
14
14a
+
-
13
DIFFERENCE IMAGE
13a
BLOCK DISCRETE COSINE TRANSFORM
15
15a
ENERGY > THRESHOLD
21
NORMALIZE
16
FREQUENCY WEIGHT
17
QUANTIZER
18
CODER
20a
CODED DIFFERENCE IMAGE
22
DIFFERENTIAL PULSE CODE MODULATION
19
CODER
20b
CODED DOWNSAMPLED IMAGE
23
10
11a

Fig. 1

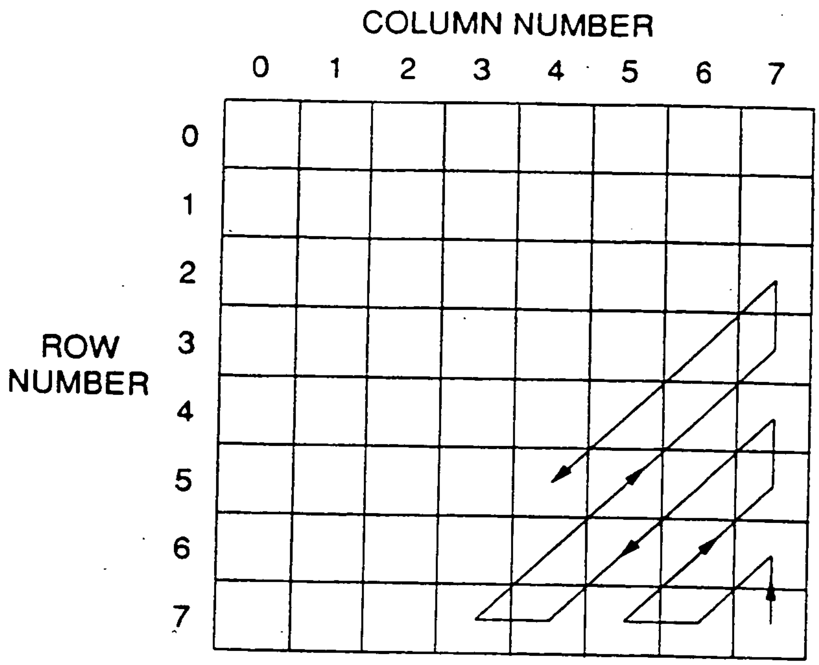

COLUMN NUMBER
0
1
2
3
4
5
6
7
ROW NUMBER
0
1
2
3
4
5
6
7

Fig. 2

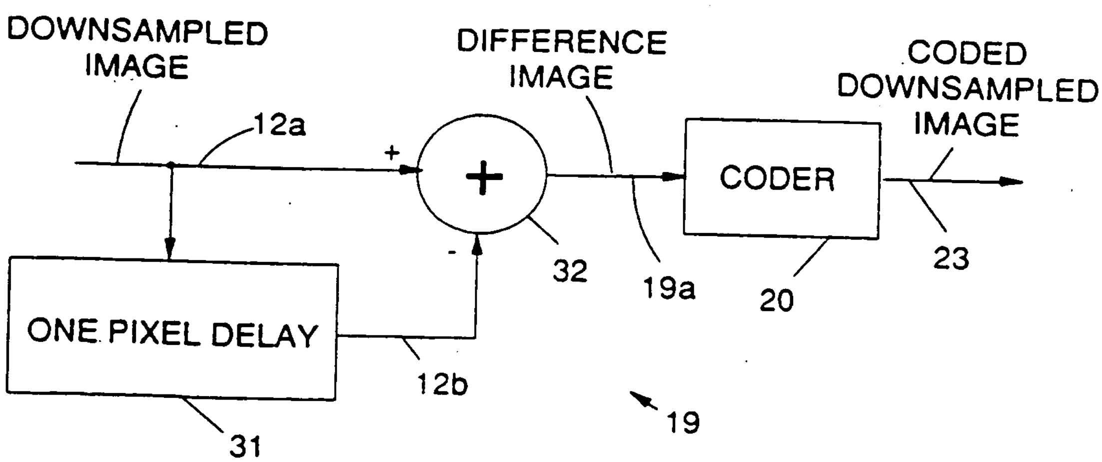

DOWNSAMPLED IMAGE
12a
+
+
-
DIFFERENCE IMAGE
CODER
CODED DOWNSAMPLED IMAGE
32
19a
20
23
ONE PIXEL DELAY
12b
19
31

Fig. 3

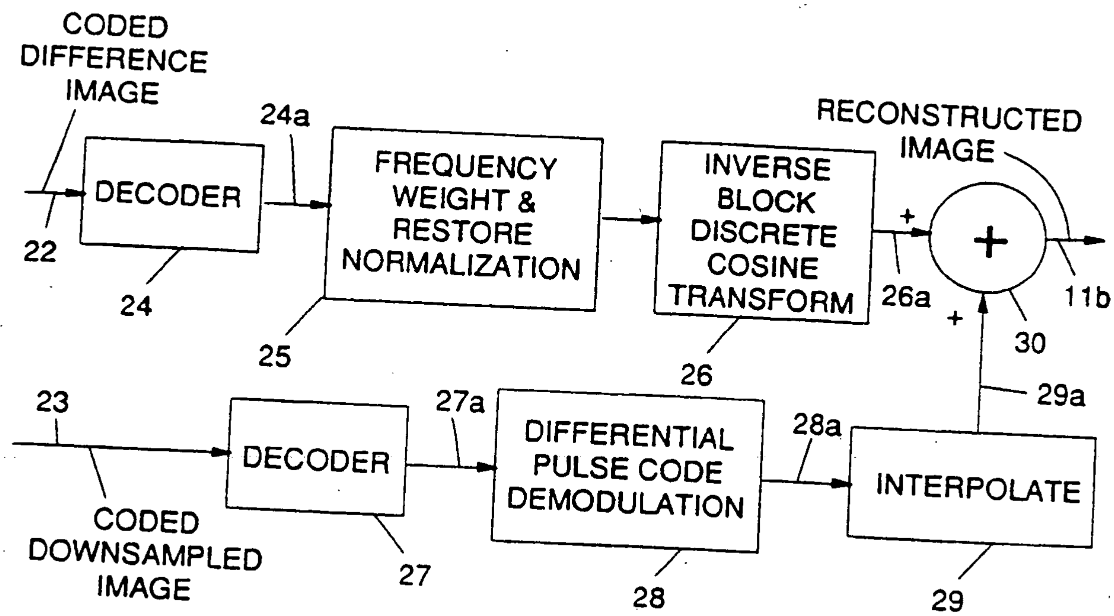

CODED
DIFFERENCE
IMAGE
22
DECODER
24
24a
FREQUENCY
WEIGHT &
RESTORE
NORMALIZATION
25
INVERSE
BLOCK
DISCRETE
COSINE
TRANSFORM
26
26a
RECONSTRUCTED
IMAGE
+
+
11b
30
23
DECODER
CODED
DOWNSAMPLED
IMAGE
27
27a
DIFFERENTIAL
PULSE CODE
DEMODULATION
28
28a
INTERPOLATE
29
29a

Fig. 4

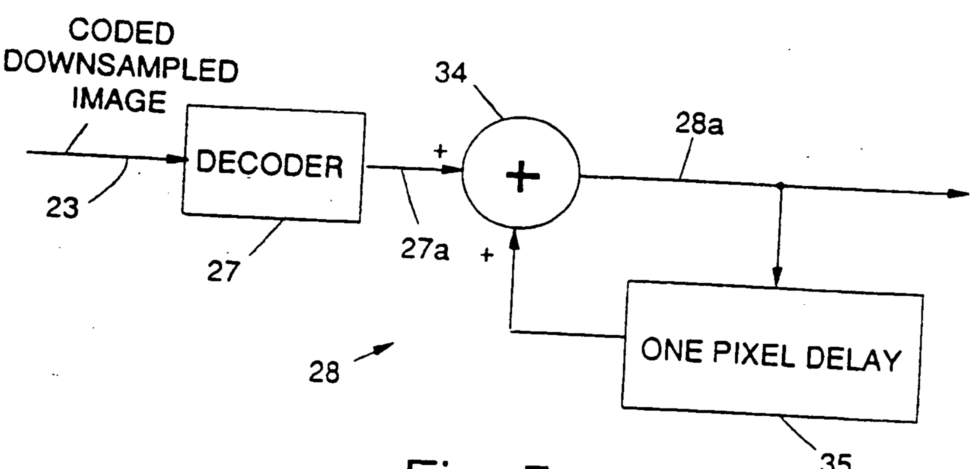

CODED
DOWNSAMPLED
IMAGE
23
DECODER
27
27a
34
+
+
28a
ONE PIXEL DELAY
35
28

Fig. 5

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 93 30 5022

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
Y	US-A-4 665 436 (OSBORNE ET AL.) * column 4, line 44 - column 7, line 27; figure 1 *	1,5,10, 11,15	H04N7/13 H04N7/133
Y	IEEE TRANSACTIONS ON COMMUNICATIONS vol. 37, no. 4, April 1989, NEW YORK US pages 380 - 386 NG ET AL. 'Two-Tier DPCM Codec for Videoconferencing' * page 380, column 1, line 21 - page 381, column 1, line 37; figure 1 *	1,5,10, 11,15	
Y	IEICE TRANSACTIONS ON COMMUNICATIONS vol. E75-B, no. 5, May 1992, TOKYO JP pages 340 - 348 ALGAZI ET AL. 'Perceptually Transparent Coding of Still Images' * page 340, column 2, line 14 - page 341, column 2, line 18; figure 1 * * page 342, column 2, line 1 - page 344, column 2, line 9; figure 2 *	1,5,10, 11,15	TECHNICAL FIELDS SEARCHED (Int. Cl.5)
A	SYSTEMS & COMPUTERS IN JAPAN vol. 22, no. 2, 1991, NEW YORK US pages 85 - 95 GOTOH ET AL. 'A Compression Algorithm for Medical Images and a Display with the Decoding Function' * page 86, column 1, line 5 - page 88, column 1, line 21; figures 1,2 *	1,10,11	H04N

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
BERLIN	16 AUGUST 1993	MATERNE A.

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P0401)

433a Electronics and Communications in Japan
73(1990)June, No.6, Part I, New York, US

H04N 7/133

G06F15/64C

Electronics and Communications in Japan, Part 1, Vol. 73, No. 6, 1990
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. 72-B-I, No. 4, April 1989, pp. 377-384.

E p. 12 - 21

H04N7/133 H

Improvement of Picture Quality and Coding Efficiency Using Discrete Cosine Transform

Yuuji Izawa and Masaaki Takizawa, Members

Central Research Laboratory, Hitachi, Ltd., Kokubunji, Japan 185

SUMMARY

In the applications of still picture coding such as image data base, the progressive coding is effective in reducing the retrieval time and the transmission cost. For the block coding based on the discrete-cosine transform (DCT), a study has been made on the progressive and successive transmission from dc to higher-order ac components. In this scheme, the major cause of the picture quality degradation is the block distortion produced at the initial stage.

This paper proposes two methods to reduce the block distortion where the lower-order ac coefficients are determined from the dc coefficients transmitted at the first stage and applied the inverse-DCT. The first method is to determine the lower-order ac coefficients so that pixel values are smoothly connected at some points on the block boundary. The second method is based on the application of the interpolation filter utilizing the property of DCT. Furthermore, it is shown that by transmitting the difference between the DCT coefficients to be transmitted and the coefficients generated by the proposed method, the coding efficiency can be improved by 5 to 8 percent.

1. Introduction

In picture coding for low bit rate, a high compression ratio is achieved by transform coding. The orthogonal transform is a method which transforms the picture signal into a sort of two-dimensional (2-D) frequency component. Among such methods, DCT (discrete-cosine transform) [1 - 4] is used extensively.

On the other hand, in the still-picture transmission, assuming the applications such as image database, the progressive coding [5, 6] is considered useful, from the viewpoint of reducing the retrieval time, transmission cost and the psychological load.

A study has been made which combines those methods to transmit and display successively, from dc coefficients to the higher-order ac coefficients. However, this method has a problem in that the block distortion is produced at the initial stage. The 2-D filtering can be applied to eliminate the block distortion. However, it causes a large computational complexity and a special hardware is required for the real-time processing.

This paper proposes two methods to reduce the block distortion by a small amount of calculation. Those methods determine the lower-order ac coefficients from dc coefficients transmitted at the first stage and then applied the inverse-DCT [7, 8]. Furthermore, the average separating block coding [9] is applied to the proposed method. By simulation, the improvement of the coding efficiency is examined and the computational complexities are compared.

2. Block Coding by DCT

This section describes the fundamental properties of the DCT.

2.1 Definition of DCT

The n-th order normalized DCT matrix $[T_{ij}]$ $(i \times j)$ is given as follows:

$$T_{ij}=\sqrt{2/n}\cdot k_i\cdot\cos\left[(i-1)\left(j-\frac{1}{2}\right)\pi/n\right] \quad (i, j=1, 2, \cdots, n) \qquad (1)$$

where $k_i=1/\sqrt{2}$ for $i = 1$
$k_i = 1$ for $i \neq 1$.

ISSN8756-6621/90/0006-

433a Electronics and Communications in Japan
73(1990)June., No.6, Part I, New York, US

i indicates the spatial frequency, and j denotes the parameter indicating the position in the block; $[T_{ij}]$ is a unitary matrix.

Letting the original image be $[G_{ij}]$ and the transform coefficient be $[C_{ij}]$, the transform and the inverse transform are defined as follows:

$$[C_{ij}]=[T_{ij}]\cdot[G_{ij}]\cdot{}^t[T_{ij}] \qquad (2)$$

$$[G_{ij}]=[T_{ij}]^{-1}\cdot[C_{ij}]\cdot\{{}^t[T_{ij}]\}^{-1} = {}^t[T_{ij}]\cdot[C_{ij}]\cdot[T_{ij}] \qquad (3)$$

2.2 Fundamental properties of DCT

The DCT is equivalent to the following procedure. By mirror image, a 2-D even function is generated from the original image, and the discrete Fourier transform (DFT) is applied. In DFT, the original image is sampled by a finite number of impulses. At the same time, the image periodicity is tacitly assumed. Then, the spectrum is equal to a line spectrum, and DFT can be replaced by a Fourier series expression.

In other words, an arbitrary 2-D even function can be expressed as a superposition of 2-D cosine functions, in which the aliasing component by sampling is eliminated. Figure 1 shows the images corresponding to the lower-order DCT coefficients C_{ij}. Using the forementioned property, an interpolation filter can be constructed. This filter and its application are discussed in Sect. 4.

2.3 Block distortion in progressive coding

Essentially, the block coding causes the block distortion. In the progressive still picture coding, the distortion near the block boundary is enhanced when the picture signal is restored from dc coefficients.

One of the general methods to eliminate the block distortion is to employ the 2-D filter. In this method, the impulse response of the ideal filter to eliminate the frequency components above one-half the sampling frequency is prepared. Then the 2-D convolution of the impulse response and the picture signal is calculated.

This method has the following features.

(1) Using the impulse response of the ideal filter, an ideal interpolation can be realized.

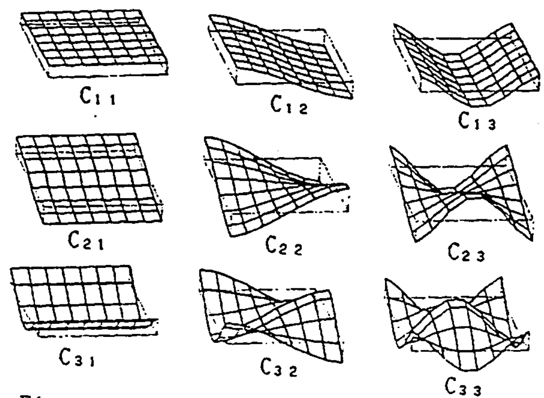


Fig. 1. Images correspond to DCT coefficients.

(2) In practice, the impulse response is truncated by a finite duration using a window and some errors are produced.

(3) The product-sum operation is performed using a 2-D operator ($m \times m$) for each pixel to be interpolated, and the computational complexity increases in proportion to the square of m.

Such a convolution can be replaced by a product form in the Fourier transform domain. Consequently, there is a possibility that the computational complexity of filtering is reduced by operating on the DCT coefficients. The implementation of this idea is discussed in the next section.

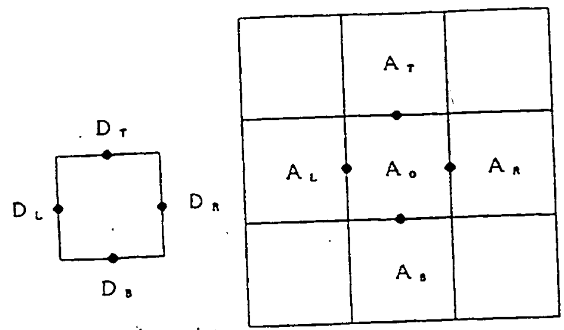
3. Elimination of Block Distortion by Manipulation of Transform Coefficient

Using the averages of adjacent blocks, the lower-order ac coefficients can be estimated. Then by the inverse DCT, the block distortion is eliminated. This method is shown as follows.

3.1 Manipulation of coefficients

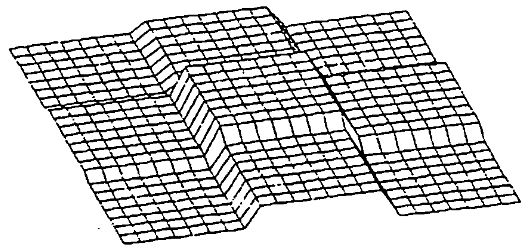
When the inverse DCT is applied to dc coefficients and N unknown ac coefficients, the pixel value in the block can be represented by a superposition of $N + 1$ cosine surfaces corresponding to those coefficients. These surfaces are continuous, as shown in Fig. 1. When values of arbitrary N points in the block (they need not necessarily be on the pixel) are given, the forementioned coefficients can be determined.

433a Electron and Communications in Japan
73(1990)June, No.6, Part I, New York, US

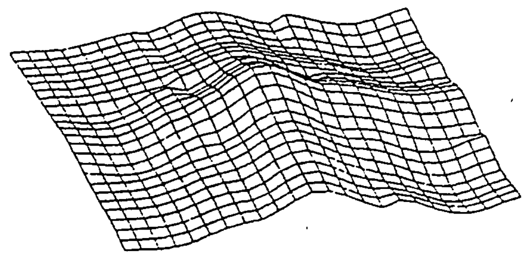


(a) Connection points on block boundary
(b) Average of blocks

Fig. 2. Four-points connection method.



(a) Original image



(b) Interpolated image

Fig. 3. Interpolated image by four-points connection method.

The procedure to estimate the coefficients smoothly connecting at the block boundary is as follows.

(1) The values at the connecting points (N in number) on the block boundary are determined from dc coefficients of adjacent blocks.

(2) The inverse DCT is applied to N unknown coefficients and the known dc coefficient, and the values on the block boundary are calculated. Then, the values determined in (1) are substituted.

(3) The equations obtained in (2) are solved to determine the values of unknown coefficients.

It is reasonable to define the points on the block boundary at the four corners or on four edges. By calculating beforehand the operations of (2) and (3) for those points, the computational complexity can substantially be reduced. The calculation examples for four and eight unknown coefficients (connections) are shown in the following.

3.2 Four-point connection

By applying the inverse DCT to the known dc coefficient C_{11} and four coefficients C_{12}, C_{13}, C_{21} and C_{31} near the dc component, the following expression is obtained:

$$\begin{aligned} D_{ij}=[&C_{11}+\sqrt{2}\cdot C_{12}\cdot\cos\{(2j-1)\pi/2n\} \\ &+\sqrt{2}\cdot C_{13}\cdot\cos\{2(2j-1)\pi/2n\} \\ &+\sqrt{2}\cdot C_{21}\cdot\cos\{(2i-1)\pi/2n\} \\ &+\sqrt{2}\cdot C_{31}\cdot\cos\{2(2i-1)\pi/2n\}]/n \\ &(i, j=1, 2, \cdots, n) \end{aligned} \tag{4}$$

where i and j are assumed as integers. However, D_{ij} can be considered as a superposition of cosine surfaces which are continuous in regard to i and j. Then, using real numbers i and j, the value at an arbitrary point in the block can be represented.

Consider, for example, the case of n = 8, and assume that four points are set on the block boundary as in Fig. 2(a). The values D_L, D_R, D_T and D_B are given as follows:

$$D_L=[C_{11}+\sqrt{2}\cdot C_{12}+\sqrt{2}\cdot C_{13}-\sqrt{2}\cdot C_{31}]/8 \tag{5}$$
$$D_R=[C_{11}-\sqrt{2}\cdot C_{12}+\sqrt{2}\cdot C_{13}-\sqrt{2}\cdot C_{31}]/8 \tag{6}$$
$$D_T=[C_{11}+\sqrt{2}\cdot C_{21}-\sqrt{2}\cdot C_{13}+\sqrt{2}\cdot C_{31}]/8 \tag{7}$$
$$D_B=[C_{11}-\sqrt{2}\cdot C_{21}-\sqrt{2}\cdot C_{13}+\sqrt{2}\cdot C_{31}]/8 \tag{8}$$

Using the averages A_0, A_L, A_R, A_T, and A_B of adjacent blocks as shown in Fig. 2(b), they can be represented as follows:

$$D_L=[A_L+A_0]/2 \tag{9}$$
$$D_R=[A_0+A_R]/2 \tag{10}$$
$$D_T=[A_T+A_0]/2 \tag{11}$$
$$D_B=[A_0+A_B]/2 \tag{12}$$

where

$$A_0=C_{11}/8 \tag{13}$$

433a Electronics and Communications in Japan
73(1990)June, No.6, Part I, New York, US

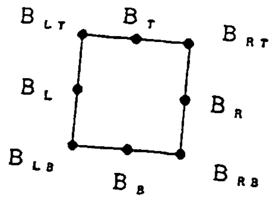


Fig. 4. Connection points on block boundary (eight-points method).

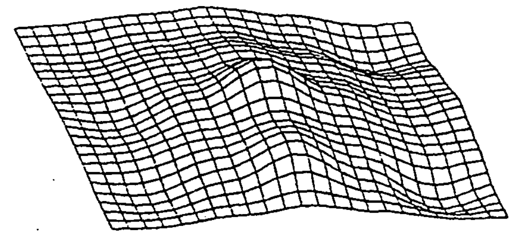


Fig. 5. Interpolated image by eight-points connection method.

Eliminating D_L, D_R, D_T and D_B from those expressions, the following equations can be obtained:

$$C_{12}=\sqrt{2}\cdot(A_L-A_R) \tag{14}$$

$$C_{13}\doteqdot\sqrt{2}\cdot(A_L+A_R-2\cdot A_0) \tag{15}$$

$$C_{21}=\sqrt{2}\cdot(A_T-A_B) \tag{16}$$

$$C_{31}\doteqdot\sqrt{2}\cdot(A_T+A_B-2\cdot A_0) \tag{17}$$

Applying the inverse DCT to those four coefficients and dc coefficient C_{11}, a smooth connection is realized on the block boundary (mid-points of edges).

Figure 3(b) is a 2-D representation of the effect of the forementioned method. Compared with the original image (a), a smoother surface is generated and the block distortion is reduced.

3.3 Eight-point connection

When the unknown coefficients are defined as 8, i.e., $C_{ij}\ (i, j=1, 2, 3 \quad i=j\neq 1)$, the same method mentioned in the foregoing can be applied. Assume, for example, that eight points on the boundary are defined as in Fig. 4. Then the following solution is obtained:

$$C=\mathrm{K}\cdot B \tag{18}$$

$$C=\begin{bmatrix} C_{12} \\ C_{13} \\ C_{21} \\ C_{22} \\ C_{23} \\ C_{31} \\ C_{32} \\ C_{33} \end{bmatrix} \qquad B=\begin{bmatrix} B_{LT}-A_0 \\ B_T-A_0 \\ B_{RT}-A_0 \\ B_{LB}-A_0 \\ B_B-A_0 \\ B_{RB}-A_0 \\ B_L-A_0 \\ B_R-A_0 \end{bmatrix} \tag{19}$$

$$\mathrm{K}=\begin{bmatrix} 1/\sqrt{2}, & 0, & -1/\sqrt{2}, & 1/\sqrt{2}, & 0, & -1/\sqrt{2}, & \sqrt{2}, & -\sqrt{2} \\ 1/\sqrt{2}, & 0, & 1/\sqrt{2}, & 1/\sqrt{2}, & 0, & 1/\sqrt{2}, & \sqrt{2}, & \sqrt{2} \\ 1/\sqrt{2}, & \sqrt{2}, & 1/\sqrt{2}, & -1/\sqrt{2}, & -\sqrt{2}, & -1/\sqrt{2}, & 0, & 0 \\ 1, & 0, & -1, & -1, & 0, & 1, & 0, & 0 \\ 1/2, & -1, & 1/2, & 1/2, & 1, & -1/2, & 0, & 0 \\ 1/\sqrt{2}, & \sqrt{2}, & 1/\sqrt{2}, & 1/\sqrt{2}, & \sqrt{2}, & 1/\sqrt{2}, & 0, & 0 \\ 1/2, & 0, & -1/2, & 1/2, & 0, & -1/2, & -1, & 1 \\ 0, & -1, & 0, & 0, & -1, & 0, & -1, & -1 \end{bmatrix} \tag{20}$$

In the foregoing, the boundary value B can be determined by the weighted sum of the average values of the adjacent blocks. The inverse DCT circuit could be employed in those matrix manipulations. Figure 5 shows the effect of the processing by the forementioned method [the original image is the same as in Fig. 3(a)]. Comparing the result with the case of four-point connection, it is seen that smooth connections are realized at four corners of the block.

4. Elimination of Block Distortion Using DCT Interpolation Filter

This section describes the image interpolation filter, which is based on the fact that the surfaces corresponding to the DCT coefficients are continuous, as well as its application to the elimination of the block distortion.

4.1 Interpolation filter

In the following, a method is described which derives the interpolated image $[Q_{ij}]$ of $N \times N$ pixels from the original image $[G_{ij}]$ of $n \times n$ pixels. As already discussed, the coefficient $[C_{ij}]$ $(i, j = 1, 2, \ldots, n)$ corresponds to a continuous surface with a particular spatial frequency component. In other words, this surface can be regarded as an interpolated image, and the value at an arbitrary point on the surface can be calculated from $[C_{ij}]$.

Assuming that the interpolation is performed by discrete processing, the interpolated image $[Q_{ij}]$ $(i, j = 1, 2, \ldots, N)$ is represented as follows:

$$[Q_{ij}] = {}^t[H_{ij}] \cdot [C_{ij}] \cdot [H_{ij}] \tag{21}$$

$[H_{ij}]$ is an $n \times N$ matrix with the following expression:

$$H_{ij} = \sqrt{2/n} \cdot K_i \cdot \cos\left[(i-1)\left(j-\frac{1}{2}\right)\pi/N\right] \quad (i=1, 2, \cdots, n)\ (j=1, 2, \cdots, N) \tag{22}$$

where $K_i = 1/\sqrt{2}$ for $i = 1$

$K_i = 1$ for $i \neq 1$.

The foregoing description is divided into two stages: 1) to derive the DCT coefficients; and 2) the generation of the interpolated image using the derived coefficients. However, we can combine these stages into a single manipulation. Using both expressions, the following relation can be derived:

$$\begin{aligned}[Q_{ij}] &= {}^t[H_{ij}] \cdot [T_{ij}] \cdot [G_{ij}] \cdot {}^t[T_{ij}] \cdot [H_{ij}] \\ &= \{{}^t[H_{ij}] \cdot [T_{ij}]\} \cdot [G_{ij}] \cdot \{{}^t[T_{ij}] \cdot [H_{ij}]\}\end{aligned} \tag{23}$$

where $\{{}^t[H_{ij}] \cdot [T_{ij}]\}$ and $\{{}^t[T_{ij}] \cdot [H_{ij}]\}$ are $(N \times n)$ and $(n \times N)$ matrices, respectively. When the positional relation of the pixels is given to the original and the interpolated images, they have constant values independent of the contents of these images. When the interpolation is not made $(N = n)$, they turn out to be unit matrix, and the original image $[G_{ij}]$ is reproduced.

In the definition of H_{ij}, j is defined as an integer. But the sampling interval need not be uniform. If the interpolation points are arranged nonuniformly, any real number can be assigned to j. Thus, the original image $[G_{ij}]$ is assumed as an $n \times n$ square block. However, by modifying the sizes of matrices $[T_{ij}]$ and $[H_{ij}]$ to be multiplied, a similar processing can be applied to the nonsquare block.

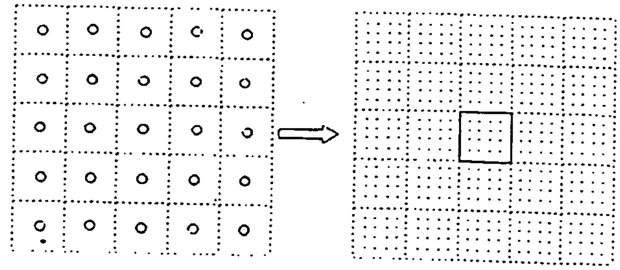


Fig. 6. Positional relation of interpolation.

4.2 Elimination of block distortion by interpolation filter

In the following, a method is shown which determines the interpolated image $[Q_{ij}]$ using only the block averages $[A_{ij}]$ instead of the original image $[G_{ij}]$. The interpolated image $[Q_{ij}]$ is affected more severely by the assumption of periodicity as one approaches the boundary. Consequently, it is advantageous to set the original image for interpolation as large as possible.

In practice, however, the value of $N(n)$ is limited due to the limitation of the size of the forementioned matrices. Then it is advisable to perform the calculation only for the central part $([Q_{ij}'])$ with a high interpolation accuracy, and to shift successively the central part.

Using this method, an adequate picture quality can be obtained. However, in the interpolation processing by assuming the block average as the pixel value of the original image, the block average of the interpolated image does not necessarily agree with the original average. This discrepancy is more remarkable near the pixels with the maximum and the minimum value.

From such a viewpoint, an interpolation method is considered in the following, where the dc value (average) of the blocks is preserved. Figure 6 shows the relation between the block average $[A_{ij}]$ and the image $[Q_{ij}]$ after interpolation processing. Assume that

433 Electronics and Communications in Japan
73(1990)June, No.6, Part I, New York, US

Table 1. Comparison of computational complexities of interpolation methods

Interpolation methods	No. of multiplications for a block	Numerical example (n = 8)		
		$K=3$	$K=4$	$K=5$
Traditional convolution	$n^2 \cdot K^2$	576	1024	1600
4-point connection	8	8		
8-point connection	$8^2(+4^3+4^2\cdot 3)$**	$64(+112)$**		
DCT	$n^2\cdot K+n\cdot K^2$	264	384	520
DCT average preserving *)	$n^2\cdot K+n\cdot K^2+K^4$	345	640	1145

*Excluding $[\mathrm{DCT}]^{-1}$. **( ) is an example of boundary-value calculation

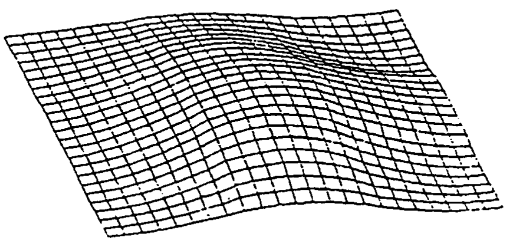


Fig. 7. Interpolated image by DCT method.

DCT is applied to the original image $[A_{ij}]$ $(i, j = 1, 2, \ldots, K)$ composed of $K \times K$ block averages. Then, $K \times K$ coefficients $[C_{ij}]$ are obtained by transform:

$$[C_{ij}]=[T_{ij}]\cdot[A_{ij}]\cdot{}^t[T_{ij}] \tag{24}$$

Assume that each block is composed of $n \times n$ pixels. The interpolated image $[Q_{ij}]$ $(i, j = 1, 2, \ldots, K\cdot n)$ composed of $(K \times n)$ pixels both in horizontal and vertical directions is derived as follows:

$$[Q_{ij}]={}^t[H_{ij}]\cdot[C_{ij}]\cdot[H_{ij}] \tag{25}$$

The following relation is formed to the average:

$$A_{ij}=\sum_{i_n, j_n} Q_{i\cdot n-n+i_n,\, j\cdot n-n+j_n} \tag{26}$$

where $(i, j=1, 2, \cdots, K)$ $(i_n, j_n=1, 2, \cdots, n)$.

By these equations, K^2 linear equations with the number equal to the unknowns $[C_{ij}]$ is generated. By solving this, the coefficients $[C_{ij}]$ can be determined. In other words, the following relations are conducted:

$$A=[W]\cdot C \tag{27}$$

$$C=[W]^{-1}\cdot A \tag{28}$$

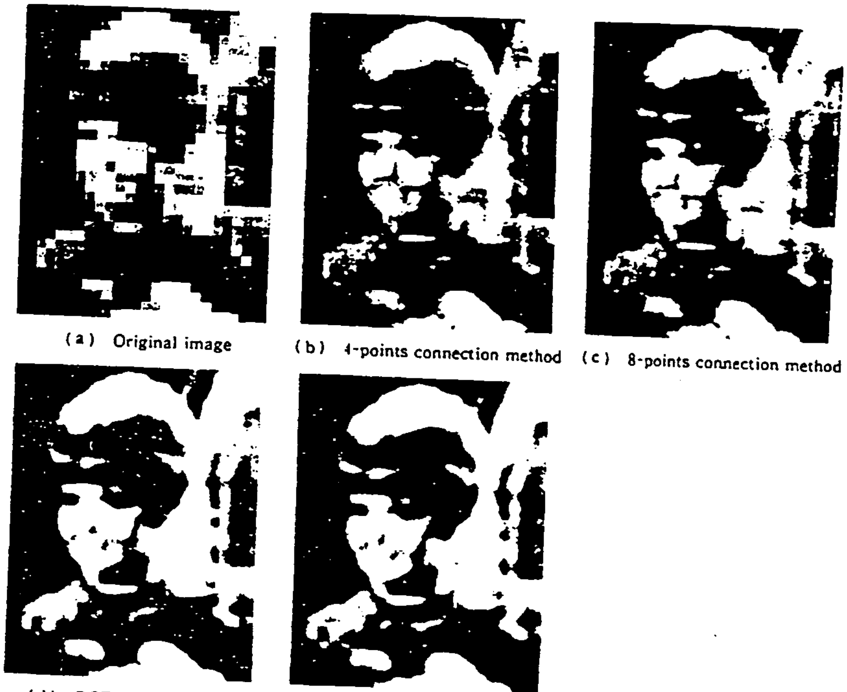
In the foregoing, C and A are vectors composed of K^2 components, which are one-dimensional rearrangements of $[C_{ij}]$ and $[A_{ij}]$, respectively. Regular matrices $[W]$ and $[W]^{-1}$ can be calculated beforehand since they do not depend on the contents of the image. This procedure can be summarized as follows.

(1) Considering the number of pixels in the block and the arrangement of blocks for the interpolation processing, matrix $[W]$ is determined. Then the inverse matrix $[W]^{-1}$ can be calculated.

(2) Vector A is generated from block averages $[A_{ij}]$, and C is determined by multiplying $[W]^{-1}$.

(3) Matrix $[C_{ij}]$ is derived from vector C. Using $[H_{ij}]$, the interpolated image $[Q_{ij}']$ $(i, j = 1, 2, \ldots, n)$ is calculated preserving the block averages.

433a Electron[ics] and Communications in Japan
73(1990)June, No.6, Part I, New York, US



(a) Original image (b) 4-points connection method (c) 8-points connection method

(d) DCT method (4 × 4) (e) DCT method (5 × 5)

Fig. 8. Examples of interpolated image.

(4) Procedures (2) and (3) are iterated blockwise.

Figure 7 illustrates the forementioned procedure in a 2-D form [original image is Fig. 3(a)]. It can be seen that a smooth cosine surface eliminating the block distortion is constructed.

4.3 Comparison of computational complexity

In the following, the computational complexities of the interpolation method are compared. For simplicity, the square block is considered. The original image is assumed as composed of $K \times K$ pixels (or blocks), and the central block as composed of $n \times n$ interpolation pixels. Then the amount of multiplications in each interpolation method is as shown in Table 1.

In four- and eight-point connections, the computational complexity is reduced by one to several, compared with the traditional convolution method, even if the number of multiplications in the inverse DCT ($3^2 \times 8 + 3 \times 8^2 = 264$) is considered. Similar merit is observed in the method by DCT.

5. Simulation Results

This section shows the simulation results for the interpolation method proposed in the previous section.

5.1 Coefficient manipulation

Figure 8(b) shows the result of the coefficient manipulation processing by four-point connection. Compared with case (a), where the block average is displayed directly, the picture quality is improved greatly. Figure 8(c) shows the result of processing by eight-point connection. Compared with the four-point connection, the distortions at four corners are reduced.

In the foregoing calculations, the boundary value B is determined using 4×4 (corner), 4×3 and 3×4 (mid-point of edge) operators as shown in Figs. 9(a), (b) and (c), respectively, for the adjacent block averages.

5.2 Interpolation filter

Using interpolation filtering, the central block (8×8 pixels) was interpolated from the 4×4 and 5×5 peripheral block

433a Electronics and Communications in Japan
73(1990)June, No.6, Part I, New York, US

-1/36	-1/36	-1/36	-1/36
-1/36	1/3	1/3	-1/36
-1/36	1/3	1/3	-1/36
-1/36	-1/36	-1/36	-1/36

(a) Corner of block

-1/30	-1/30	-1/30	-1/30
-1/30	2/3	2/3	-1/30
-1/30	-1/30	-1/30	-1/30

(b) Center of vertical side

-1/30	-1/30	-1/30
-1/30	2/3	-1/30
-1/30	2/3	-1/30
-1/30	-1/30	-1/30

(c) Center of horizontal side

Fig. 9. Example of operators to calculate the value of block boundary.

averages. The results are shown in Figs. 8(d) and (e). Although a smoother picture than the coefficient manipulation is obtained, some distortion is observed near the block boundary for the 4 × 4 case. The 4 × 4 operator is considered to be insufficient to secure the continuity between adjacent blocks. When extended to 5 × 5, the distortion is reduced sufficiently.

It is seen also that the distortions inherent to the coefficient manipulation, such as the step near the block corner in four-point connection and the unevenness near the block center in eight-point connection, are eliminated almost completely. The interpolation filter not preserving the block average described in Sect. 4.1 is also applied and verified that nearly the same picture quality is obtained.

6. Improvement of Coding Efficiency

In orthogonal transform, some correlations remain among the transform coefficients of the adjacent blocks. In the natural image, on the other hand, the probability is very low that there exists a steep edge at the block boundary, which implies that the pixels on both sides of the block boundary have a high correlation. Thus, there are some correlations between the lower-order ac coefficients including the dc component of adjacent blocks. Consequently, the coding efficiency can be improved by utilizing those correlations.

6.1 Average separating block coding

The average separating block coding is described in [9]. This method which improves the coding efficiency by transmitting the difference between the original image and the interpolated image generated by the average of adjacent blocks. Using the interpolated image generated by the DCT coefficient manipulation or by the interpolation filtering described in Sect. 3.4, the coding efficiency can be improved by a simple operation (i.e., subtraction) of the transform coefficients.

6.2 Mean square error

The mean-square error is used as an index for evaluating the coding efficiency. In the following, the value obtained as a mean square of the difference between the original image and the interpolated images is compared.

Table 2 shows the result of actual measurement for a standard picture. The improvement of coding efficiency was 5 to 6 percent by coefficient manipulation, and 8 percent by the DCT interpolation filtering. In comparison with the method based on quadratic surface described in [9], the efficiency is decreased by 1 to 2 percent in the coefficient manipulation. This is due to the effect of limiting the number of transform coefficients to 4 or 8. The computational complexity is less than one-eighth compared with 1433 in the quadratic surface method.

7. Conclusions

This paper discussed two methods for eliminating block distortion in the block coding based on the DCT. The methods determine the lower-order ac coefficients from dc coefficients, and applies the inverse DCT. The two methods are as follows.

Table 2. Comparison of mean-square error

	MSE pixel	Ratio	Coding efficiency
Block average	190.4	100.0%	± 0 %
4-point connection	147.1	77.2%	+ 5 %
8-point connection	139.1	73.1%	+ 6 %
Quadratic surface [9]	135.9	71.4%	+ 7 %
DCT interpolation*	128.9	67.7%	+ 8 %

SIDBA GIRL (256 × 256 pixel), *K = 4, n = 8.

(1) The lower-order ac coefficients are manipulated to realize a smooth connection at the block boundary.

(2) Using the interpolation filter based on the property of the DCT, ac coefficients are determined so that a smooth cosine surface is constructed.

By simulation, the improvement of the picture quality was verified. As a further application, the method to transmit the difference between the coefficient to transmit and the coefficient generated from dc components is applied (average separation block coding). It was then verified that the coding efficiency can be improved by 5 to 8 percent, compared with the traditional method. The elimination of the block distortion described in the first half of this paper can also be applied to the interframe prediction of the moving picture. One of the topics for future study is the application of the method to motion compensation.

Acknowledgement. The authors are indebted for the guidance in this study to Dr. T. Fukinuki, Research Organizer, Dr. T. Tanaka, Head of 5th Div. of Lab., and Dr. M. Nakamura, Director of R & D Center, Tech. Man. Div.

REFERENCES

1. T. Fukinuki. Digital Signal Processing of Images. Nikkan Kogyo Press (1981).
2. N. Ahmed, T. Natarajan, and K. R. Rao. Discrete-cosine transform. IEEE Trans. Comput., C-23, pp. 90-93 (Jan. 1974).
3. W. H. Chen et al. A fast computational algorithm for the discrete cosine transform. IEEE Trans. Commun., COM-25, pp. 1004-1009 (Sept. 1977).
4. W. H. Chen et al. Adaptive coding of monochrome and color images. IEEE Trans. Commun., COM-25, pp. 1285-1292 (Nov. 1977).
5. Y. Yasuda, M. Takagi, S. Kato, and T. Awano. Stepwise transmission and display of still pictures by progressive coding. Trans. (B), I.E.C.E., Japan, J63-B, 4, pp. 379-386 (Apr. 1980).
6. S. Kim, S. Kato, and Y. Yasuda. Expansion-contraction of picture suited to hierarchical transmission. J. Image Elect., 17, 2, pp. 54-60 (Apr. 1988).
7. Y. Izawa, J. Kimura, and M. Takizawa. A study of picture quality improvement in progressive still picture coding. Picture Coding Symp., Japan, 87, 4-1 (1987).
8. Y. Izawa, J. Kimura, and M. Takizawa. Elimination of block distortion in progressive still picture coding. Nat. Conv. Spr., I.E.I.C.E., Japan, D-113 (1988).
9. Y. Kato. Improvement of coding efficiency in average separating block coding. Nat. Conv. Gen., I.E.I.C.E., Japan, 1158 (1987).

433a Electronics and Communications in Japan
73(1990)June, No.6, Part I, New York, US

AUTHORS (from left to right)

Yuuji Izawa graduated in 1976 from the Dept. Indust. Mech., Fac. Eng., University of Tokyo, and obtained a Master's degree from there in 1978. He then affiliated with Hitachi, Ltd. In the Central Research Lab., he is engaged in research on switched capacitor filter and digital signal processing.

Masaaki Takizawa graduated in 1975 from the Dept. Electrical Eng., Fac. Eng., University of Tokyo, and obtained a Master's degree from there in 1977. He then affiliated with Hitachi Central Research Lab. He is engaged in research on highly efficient coding methods such as high-speed facsimile and TV codec. He received the Encour. Prize from I.E.I.C.E., Japan, in 1982.